

CAPÍTULO 5

CONCLUSIONES

5. CONCLUSIONES

Después del análisis realizado fue posible entender el comportamiento de los errores a lo largo de los diferentes tipos de procedimientos de comparación múltiple. Esto permite clarificar la idea de que el cálculo de los errores en las comparaciones dependientes e independientes son procedimientos diferentes que muestran comportamientos distintos y por lo tanto requieren ser calculados y controlados de diferente manera, a pesar de que al inicio del presente estudio se pensó que posiblemente los errores dependientes se podrían comportar de una manera similar a los independientes, los resultados de tablas 2.1 y 4.1, muestran que esto no es así.

En el capítulo 4 se encuentran los resultados obtenidos para las pruebas planteadas en la metodología o capítulo 3, aplicadas a un experimento completamente al azar con cinco tratamientos y cuatro repeticiones por tratamiento. En el presente capítulo se muestra un resumen general de las conclusiones obtenidas.

En la sección 4.3 referente a resultados y análisis de la estimación del FWE, PCE y PFE, se muestra la tabla 4.1 que resume los resultados de dichos errores, al observar la columna del PCE y compararla con la tabla 4.3, que muestra los errores en las comparaciones particulares antes de observar los datos, se comprueba que ambos resultados son los mismos por lo que se recomienda descartar el uso del PCE para controlar los errores en las pruebas de comparación múltiple dependientes debido a que forzosamente se requiere observar y ordenar los datos para llevar a cabo el experimento; al mismo tiempo el PFE no resulta un buen controlador del error debido a que su valor es igual a una constante por el PCE. Así pues, el PCE será usado como un indicador del rechazo en todas las comparaciones cuando no se han observado los datos.

De los tres errores mencionados por Hochberg y Tamhane (1987), el FWE se muestra como el único error factible que puede ser considerado al momento de analizar una

familia, resultando las pruebas mencionadas en la tabla 4.4, referente a pruebas recomendadas para el control del FWE, como aquellas que se comportan adecuadamente con respecto a este error, siempre y cuando los datos no hayan sido observados. Las pruebas recomendadas para el caso protegido son la DMS, Tukey, SNK, Duncan, D.E, Bonferroni y Scheffé, y entre las no protegidas se encuentran Tukey, Bonferroni y Scheffé.

En los textos usuales como Montgomery (2002), Guitiérrez y De la Vara (2003), la prueba de Tukey se recomienda de una manera protegida, los resultados del presente trabajo contravienen esta recomendación ya que, según los valores arrojados por esta investigación la prueba de Tukey debiera ser recomendada no protegida, ya que el FWE de ambas pruebas ya sea de uno o dos pasos es de 0.05. Para el caso de la prueba DMS, el resultado arrojado para el FWE de la prueba protegida resulta ser el deseado de 0.05, por lo que dicha prueba demuestra para este caso tener un adecuado comportamiento; caso muy similar ocurre con la prueba de Duncan, ya que a pesar de que el valor obtenido para el caso protegido en el FWE es de 0.49, bastante aproximado al deseado, su comportamiento demuestra que la prueba debe aplicarse en dos pasos.

Las pruebas de Bonferroni y Scheffé tanto para el caso protegido como no protegido muestran resultados para el FWE de 0.033 y 0.023 respectivamente, estos valores se encuentran muy por debajo del 0.05 deseado al inicio del experimento lo que significa que ambas pruebas demuestran un comportamiento muy restrictivo y poco deseable, por tal motivo se recomienda no utilizar estas pruebas. Por otro lado la prueba D.E muestra un comportamiento interesante ya que para el caso protegido tiene un FWE de 0.042 y un PCE de 0.020, recuerde que el PCE es el tamaño de la prueba para cada una de las comparaciones dentro de la familia sin haber observado los datos, por lo que esto es un indicativo de que la prueba D.E es más hábil para detectar diferencias de medias cercanas, tratándose de pruebas de dos pasos.

Una vez que los datos han sido observados y al mismo tiempo ordenados, forzosamente es necesario utilizar el concepto comparaciones particulares para el cálculo de los errores, los resultados arrojados para el experimento planteado en esta investigación se muestran en la tabla 4.2 referente a la sección 4.4, resultados y análisis en comparaciones particulares, la investigación en este punto fue dividida en dos puntos diferentes:

- Si se consideran todas las pruebas mostradas en la tabla 4.2, resulta notorio el buen comportamiento de la prueba D.E No Protegida, debido a que resulta ser la única que reparte el error deseado de una manera homogénea a lo largo de todas las comparaciones, respetando de buena manera el valor de α establecido inicialmente. Lo cual puede ser comprobando ya que maneja un rango de [0.05230, 0.04630] que demuestra ser cerrado en comparación al resto de las pruebas como la DMS que va de [0.25843, 0.00096], Duncan de [0.18621, 0.00100] o Tukey [0.04964, 0.00004] que muestran rangos bastante amplios.

- Si solamente son consideradas las pruebas recomendadas en la tabla 4.4, que son aquellas pruebas que se comportan adecuadamente con respecto al FWE, la prueba que demuestra tener un mejor desempeño es la prueba D.E Protegida, debido a que resulta la menos restrictiva de todas las pruebas. Esto puede ser observado de una manera más clara en los resultados obtenidos en la comparación entre medias cercanas como lo son la media uno Vs la media dos, en donde el resultado para esta comparación en la prueba D.E protegida es de 0.01400 mientras que para el resto de las pruebas de dos pasos se tienen resultados para DMS de 0.00465, SNK de 0.00481, Duncan de 0.00465 y Tukey para ambos casos de 0.00038, comportamiento que se repite a lo largo de cualquier comparación contigua, demostrando que la prueba D.E resulta más hábil para detectar diferencias entre medias cercanas y mantiene al tamaño deseado para medias lejanas.

5.1 Recomendaciones para futuras investigaciones.

Partiendo de los resultados e inquietudes arrojados por la investigación, se sugieren algunos problemas que podrían ser de interés para investigadores en el área.

5.1.1 Cálculo y Control de Errores en el Caso No Paramétrico.

Existen pruebas de rango múltiple no paramétricas, como Kruskal-Wallis y Friedman, las cuales se realizan a través de cotas de Bonferroni. Es posible que los algoritmos del presente trabajo puedan generalizarse para pruebas exactas en este caso.

5.1.2 Combinación de Pruebas.

Al observar los resultados obtenidos en el capítulo 4, es posible identificar que algunas pruebas para casos específicos se comportan de una manera exacta pero para otras comparaciones su comportamiento se vuelve restrictivo; por lo cual se cree sea posible utilizar una prueba que resulte en una combinación de varias pruebas existentes y que permita adaptar la PRM adecuada a la comparación que se pretenda realizar, permitiendo al experimentador no encerrarse en una misma prueba a lo largo de todas las comparaciones. Obviamente dicha investigación requerirá de un importante análisis y estudio.

5.1.3 Generalizar a otros tratamientos y repeticiones.

Por cuestiones prácticas en la presente investigación se utilizó un mismo ejemplo aplicado a los diferentes posibles casos y errores. Se recomienda para ampliar dicho estudio generalizar el número de tratamientos y repeticiones para observar el comportamiento de los errores en más casos, esperando únicamente sustentar la robustez de esta tesis.